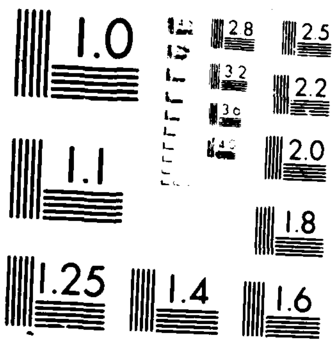
AD-A186 016 STOCHASTIC FILTERING SOLUTIONS FOR ILL-POSED LINEAR PROBLEMS AND THEIR EX..(U) NORTH CAROLINA UNIV AT CHAPEL HILL CENTER FOR STOCHASTIC PROC.. R BRIGOLA 1/1

UNCLASSIFIED MAR 87 TR-180 AFOSR-TR-87-1100 F/G 12/3 NL


1.0
2.8
2.5
3.2
2.2
2.0
1.1
1.8
1.25
1.4
1.6

DTIC FILE COPY

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

AD-A186 016

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION: UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY: NA	3. DISTRIBUTION/AVAILABILITY OF REPORT: Approved for public release; Distribution Unlimited
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE: NA	
4. PERFORMING ORGANIZATION REPORT NUMBER(S): Technical Report No. 180	5. MONITORING ORGANIZATION REPORT NUMBER(S): AFOSR-TR- 87-1100
6a. NAME OF PERFORMING ORGANIZATION: University of North Carolina; 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION: AFOSR/NM
6c. ADDRESS (City, State and ZIP Code): Statistics Dept. Phillips Hall 039-A Chapel Hill, NC 27514	7b. ADDRESS (City, State and ZIP Code): Building 410 Bolling AFB, DC 20332-6448
8a. NAME OF FUNDING/SPONSORING ORGANIZATION: AFOSR; 8b. OFFICE SYMBOL (If applicable): NM	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER: F49620 ~~85XEXEXXX~~ 82 C 0009
8c. ADDRESS (City, State and ZIP Code): Building 410, Bolling AFB, DC 20332-6448	10. SOURCE OF FUNDING NOS. PROGRAM ELEMENT NO.: 6.1102F; PROJECT NO.: 2304; TASK NO.: A5; WORK UNIT NO.

11. TITLE (Include Security Classification): Stochastic filtering solutions for ill-posed linear problems and their extensions to measurable transformations

12. PERSONAL AUTHOR(S): Brigola, R.

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Yr., Mo., Day)	15. PAGE COUNT
Preprint	FROM 9/84 to 9/86	March 1987	19

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB. GR.	

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

An ill-posed linear problem Ax=y in Hilbert space is considered as a filtering problem AX+Z=Y for Hilbert space valued random elements. Depending on the models for the signal X and the noise Z, the solutions of this problem are discussed in the context of cylinder measures on Hilbert spaces and their radonification by the Abstract Wiener space concept. Extensions of the solutions to measurable transformations are given explicitly. The filtering solution is related to the solution of the problem Ax=y obtained by Tichonov's regularization method.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT: UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS ☐	21. ABSTRACT SECURITY CLASSIFICATION: UNCLASSIFIED
22a. NAME OF RESPONSIBLE INDIVIDUAL: ~~Peggy Ravitch~~ Maj Crowley	22b. TELEPHONE NUMBER (Include Area Code): ~~919-962-2307~~ 767-5025; 22c. OFFICE SYMBOL: AFOSR/NM

DTIC ELECTE OCT 07 1987 S E D

DD FORM 1473, 83 APR EDITION OF 1 JAN 73 IS OBSOLETE. UNCLASSIFIED

AFOSR-TR- 87-1100

CENTER FOR STOCHASTIC PROCESSES

Department of Statistics
University of North Carolina
Chapel Hill, North Carolina

STOCHASTIC FILTERING SOLUTIONS FOR ILL-POSED LINEAR PROBLEMS

AND THEIR EXTENSION TO MEASURABLE TRANSFORMATIONS

by

R. Brigola

Technical Report No. 180

March 1987

STOCHASTIC FILTERING SOLUTIONS FOR ILL-POSED LINEAR PROBLEMS AND THEIR EXTENSION TO MEASURABLE TRANSFORMATIONS

by

R. Brigola
University of Regensburg
Fed. Rep. of Germany
and
Center for Stochastic Processes
University of North Carolina
Chapel Hill, NC

Abstract

An ill-posed linear problem $Ax = y$ in Hilbert space is considered as a filtering problem $AX + Z = Y$ for Hilbert space valued random elements. Depending on the models for the signal X and the noise Z, the solutions of this problem are discussed in the context of cylinder measures on Hilbert spaces and their radonification by the Abstract Wiener space concept. Extensions of the solutions to measurable transformations are given explicitly. The filtering solution is related to the solution of the problem $Ax = y$ obtained by Tichonov's regularization method.

This research has been supported by AFOSR Contract No. F 49620 82 C 0009

1. Introduction

Let H_1 and H_2 be real, separable Hilbert spaces and $A: H_1 \longrightarrow H_2$ be a linear bounded operator. By Hadamard's definition, a linear problem $Ax = y$ is well-posed if the solution exists, is unique and depends continuously on the data; otherwise a problem is called ill-posed.

Examples: 1) Fredholm integral equations of the first kind:
If $\Omega \subset \mathbb{R}^n$ is a bounded region, $k \in L^2(\Omega^2)$, $f \in L^2(\Omega)$, then

$$Kf(x) := \int_\Omega k(x,y) f(y) dy$$

is Hilbert-Schmidt in $L^2(\Omega)$.
Hence the linear problem $Kf = g$ ($g \in L^2(\Omega)$ given) is ill-posed.

2) A linear equation $Ax = y$ in $\mathbb{R}^n$ may be numerically ill-posed, if $\det A << 1$

In the following we will consider such problems from a statistical point of view as introduced in the work of O.N. Strand and E.R. Westwater [11], J.N. Franklin [2] and A. Uhlig [13]. This point of view is motivated by the following reasons:
Many methods for the calculation of unknown states x in physical or technical problems do not allow to observe the interesting state x directly, but give an observation y, whose functional relationship with x may be described by a linear equation $Ax = y$ as above. Such observations often may be affected with a random additive noise. Also the unknown state x may depend itself on a random law. Thus a linear inverse problem $Ax = y$ often is already an approximation for an equation of the form $Ax + z = y$, where x,y,z are random elements, i.e. we have a filtering problem, namely to estimate the unknown state x given a noisy observation y.
Of course, one will need additional information to solve this estimation problem, assumptions on the laws of the random elements and a loss function to optimize the estimate. However, the deterministic methods to give an approximate solution for an ill-posed problem $Ax = y$ in Hilbert space, for instance the Tichonov regularization (cf. [12]), also need additional information. The Tichonov regularization, i.e. the solution of the variational problem $\|Ax-y\|^2 + \alpha F(x) = \min!$, requires information on the smoothness of x to choose the regularization parameter α and the regularization functional F.

To give estimates of the approximation error, this method also needs to know a compact set containing the solution x.
Considering the problem as a stochastic filtering problem, those assumptions are replaced by statistical assumptions on the signal and the noise.
Let $(\Omega,\mathcal{F},P)$ be a probability space, H_1 and H_2 real, separable Hilbert spaces, X an H_1-valued random variable on $(\Omega,\mathcal{F},P)$, and Y,Z H_2-valued random variables resp. . For a linear bounded operator $A:H_1 \longrightarrow H_2$ we consider the estimation problem $AX + Z = Y$, i.e. we want to give an estimate x^* for $X(\omega)$ given $Y(\omega) = y$.
To make the paper self-contained, in the following we shortly summarize the well-known solution for finite-dimensional state spaces H_1, H_2. In section 2 we will consider the problem for infinite-dimensional state spaces.

Finite-dimensional state spaces H_1, H_2

Notations: i) $E(\langle X,h\rangle) = \langle x',h\rangle$ $(h \in H_1)$ for a suitable $x' \in H_1$;
x' is called the mean E(X) of X.
ii) $E(\langle X-x',h_1\rangle\langle X-x',h_2\rangle) = \langle Rh_1,h_2\rangle$ $\quad (h_1,h_2 \in H_1)$ with R self-adjoint, $R \geq 0$, and for X centered, (e_k) CONS in H_1:

$$E(\|X\|^2) = \sum_{k=1}^{\dim H_1} \langle Re_k, e_k\rangle = \mathrm{Tr}\, R.$$

R is called covariance of X.

Assumptions: i) X and Z independent

ii) X centered with covariance R

iii) Z centered with covariance S

We look for a linear least squares estimate $L_0:H_2 \longrightarrow H_1$, i.e.

(1) $E(\|X-L_0Y\|^2) = \min\{E(\|X-LY\|^2): L:H_2 \longrightarrow H_1 \text{ linear, bounded}\}$

Let $Q:H_2 \longrightarrow H_1$ be the cross-correlation between signal X and observation Y, i.e. $\langle h_1,Qh_2\rangle = E(\langle X,h_1\rangle\langle Y,h_2\rangle)$, $(h_1 \in H_1,\ h_2 \in H_2)$.

Calculating the error covariance:

(2) $E(\|X-LY\|^2) = Tr.[R+LKL^*-LQ^*-QL^*] =: F(L)$

where K is the covariance of Y and L^* denotes the adjoint operator of L, we see that F is a quadratic form on the Hilbert space $L(H_2,H_1)$ of linear operators from H_2 into H_1 endowed with the Hilbert-Schmidt norm

$$\|L\|_2^2 = \sum_{k=1}^{\dim H_2} <Lf_k, f_k> \qquad ((f_k)\ \text{CONS in } H_2)$$

Choosing L_0 such that $L_0K = Q$, one obtains $F(L) \geq F(L_0) = Tr.[R-L_0KL_0^*]$ $(L \in L(H_2,H_1))$.

Thus L_0 solves the estimation problem. If X,Z are Gaussian, one explicitly has $Q = RA^*$, $K = ARA^* + S$, and therefore $L_0 = RA^*(ARA^*+S)^{-1}$; here K^{-1} denotes the inverse of K if K is invertible, otherwise K^{-1} means the left-pseudo-inverse $(K^*K)^{-1}K^*$ of K.

For a given right side $y = Y(\omega)$ one thus has as LLS-estimate (linear least squares estimate)

(3) $x^* = RA^*(ARA^*+S)^{-1}y = L_0y$

Remarks

i) If X,Z are not centered, $x^* = E(X) + L_0(y-A(E(X)-E(Z)))$ is LLS-estimate

ii) If X,Z are independent Gaussian with regular covariances, then $L_0Y = E(X|Y)$, the conditional expectation of X given Y, and x^* is the mean of the conditional distribution of X given $Y(\omega) = y$

iii) L_0 minimizes the error of linear functionals of the estimate, i.e.

(4) $E(<X-L_0Y,h>^2) = \min\{E(<X-LY,h>^2) : L \in L(H_2,H_1)\} \quad (h \in H_1)$

iv) If A is invertible and $S=0$ (no noise), then $x^* = A^{-1}y$, the deterministic solution of the inverse problem, for arbitrary covariance R.

v) A connection with Tichonov's regularization method is given by the following observation:
The functional $\|Ax-y\|^2 + \sigma<R^{-1}x,x>$ attains its minimum at $x^* = RA^*(ARA^*+\sigma^2 I)^{-1}y$ (cf.[12]).

Thus Tichonov's method with regularization parameter σ and regularization functional $<R^{-1}.,.>$ gives the LLS-estimate under the assumptions of a centered signal X with regular covariance R and a centered white noise Z with covariance $\sigma^2 I$, I identity operator on H_2.

2. LLS-estimates in infinite-dimensional state spaces.

In the following let H_1 and H_2 be real, separable, infinite-dimensional Hilbert spaces. Of course, the problem in generalizing the results to infinite-dimensional spaces depends on the choice of the mathematical model for the signal X and the noise Z. The following considerations may clear the fundamentals and show the significance of the development of a finitely additive filtering theory in the work of A.V. Balakrishnan [1] and in a series of papers of G. Kallianpur and R. Karandikar (cf. [8]).

2.1 Gaussian signal and noise with nuclear covariances.

For convenience , let $A: H_1 \longrightarrow H_2$ be onto, but not continuously invertible. Let X, Z be zero-mean Gaussian, independent, with nuclear covariances R and S resp., and $Ker(ARA^* + S) = \{0\}$. By $\mathfrak{L}(H_2)$ we denote the Borel measurable subsets of H_2.
The problem is again to give an estimate x^* given $Y(\omega) = y$, where $AX + Z = Y$.
By assumption, the operator $(ARA^* + S)$ has dense range in H_2; hence it exists a unique left-inverse $(ARA^* + S)^{-1}$ with domain of definition $rg(ARA^* + S)$; but it is unbounded as inverse of a nuclear operator. Moreover, the well-known crucial point is, that $rg(ARA^* + S)$ is a set of measure zero with respect to the distribution of Y:

$$(5) \qquad P \circ Y^{-1}(rg(ARA^* + S)) = 0$$

<u>Proof.</u> Denote $ARA^* + S =: K$; then $rg(K) \subset rg(K^{\frac{1}{2}})$. If (e_k) is a CONS of eigenvectors of K with corresponding eigenvalues (λ_k), then $x \in rg(K^{\frac{1}{2}})$ iff $(\lambda_k^{-\frac{1}{2}} < x, e_k >) \in l_2(\mathbb{N})$.
But $\lambda_k^{-\frac{1}{2}} < ., e_k >$ is standard Gaussian on $(H_2, \mathfrak{L}(H_2), P \circ Y^{-1})$; hence

$$\sum_{k=1}^{\infty} \frac{< . , e_k >^2}{\lambda_k} \text{ is } P \circ Y^{-1} \text{ - a.e. divergent, i.e. } P \circ Y^{-1}(rg(K)) = 0$$

#

Thus, in this case, the estimate $x^* = L_0 y$, $L_0 = RA^*(ARA^* + S)^{-1}$, as given in [2], is useless, since it works only for observations y, which $P \circ Y^{-1}$ - a. s. do never appear.

In the following, it is shown that there exists an extension of L_0 to an operator L whose domain of definition has $P \circ Y^{-1}$ - measure one, i.e. L is a measurable transformation, with respect to $P \circ Y^{-1}$.

<u>Proposition 1.</u> Let $(e_k)_1^\infty$ be a CONS of eigenvectors of K with corresponding eigenvalues $(\lambda_k)_1^\infty$.
Let $(r_k)_1^\infty$ be a CONS of eigenvectors of R with corresponding eigenvalues $(\mu_k)_1^\infty$.
Then it holds for every $s \in rg(K)$:

$$L_0 s = \sum_{j=1}^{\infty} \mu_j^{\frac{1}{2}} \left[\sum_{i=1}^{\infty} \frac{<s, e_i>}{\lambda_i} <R^{\frac{1}{2}} A^* e_i, r_j> \right] r_j \tag{6}$$

Defining L by (6) for all $y \in H_2$ satisfying

$$\sum_{j=1}^{\infty} \mu_j \left[\sum_{i=1}^{\infty} \frac{<y, e_i>}{\lambda_i} <R^{\frac{1}{2}} A^* e_i, r_j> \right]^2 < \infty,$$

one obtains:

i) L extends L_0

ii) L is measurable and $P \circ Y^{-1}(D(L)) = 1$,
where D(L) is the domain of definition of L.

iii) $x^* := Ly$ is the mean of the conditional distribution of X given $y \in D(L)$

<u>Proof.</u> Clearly, L extends L_0, the domain D(L) of L contains $D(L_0)$, and L is $\mathfrak{L}(H_2)$-$\mathfrak{L}(H_1)$ - measurable as a limit of finite sums of measurable mappings. One has to show $P \circ Y^{-1}(D(L)) = 1$

To prove this, the following facts will be used (cf. Gihman-Skorohod [3]):

a) A well-known theorem of Kolmogorov states: if $(\zeta_k)_1^\infty$ are independent, $\mathbb{R}$-valued r.v. such that $E\zeta_k = 0$ $(k \in \mathbb{N})$ and

$$\sum_{k=1}^{\infty} E|\zeta_k|^2 < \infty \text{, then } \sum_{k=1}^{\infty} \zeta_k < \infty \text{ a.e.}$$

b) If $(\zeta_k)_1^\infty$ are $\mathbb{R}$-valued such that

$$\sum_{k=1}^{\infty} E|\zeta_k| < \infty \text{, then } \sum_{k=1}^{\infty} |\zeta_k| < \infty \text{ a.e.}$$

c) With the above notions

$$Y = \sum_{i=1}^{\infty} \lambda_i^{\frac{1}{2}} < Y, e_i > e_i \text{ and } Y_i := \frac{< Y, e_i >}{\lambda_i^{\frac{1}{2}}}$$

are standard Gaussian $(i \in \mathbb{N})$

d) $(ARA^* + S)^{-\frac{1}{2}} e_i = \lambda_i^{-\frac{1}{2}} e_i$ $(i \in \mathbb{N})$

e) By d) Ls is transformed to

$$Ls = \sum_{j=1}^{\infty} \mu_j^{\frac{1}{2}} \left[\sum_{i=1}^{\infty} < Ce_i, r_j > \frac{< s, e_i >}{\lambda_i^{\frac{1}{2}}} \right] r_j$$

where $C := R^{\frac{1}{2}} A^* (ARA^* + S)^{-\frac{1}{2}}$; consequently

$$LY = \sum_{j=1}^{\infty} \mu_j^{\frac{1}{2}} \left[\sum_{i=1}^{\infty} < Ce_i, r_j > Y_i \right] r_j$$

f) For C in e) it holds $\|C\| \le 1$, because:

$v \in D(C) = rg(ARA^* + S)^{\frac{1}{2}} \Rightarrow \quad v = (ARA^* + S)^{\frac{1}{2}} u$ for suitable $u \in H_2$,

and since $S \ge 0$

$\|Cv\|^2 = \|R^{\frac{1}{2}} A^* u\| \le < (ARA^* + S)u, u > = \|(ARA^* + S)^{\frac{1}{2}} u\|^2 = \|v\|^2$.

Thus, C has a continuous extension to all of H_2, say C_e, since D(C) is dense in H_2. Hence, $C^* = C_e^*$ exists and $\|C^*\| \le 1$; also $C^{**} = C_e$ exists and $\|C^{**}\| \le 1$. It holds $C^* = (ARA^* + S)^{-\frac{1}{2}} A R^{\frac{1}{2}}$

Now, set

$$\zeta_j := \sum_{i=1}^{\infty} <Ce_i, r_j> Y_i$$

Then for all $j \in \mathbb{N}$, $\zeta_j < \infty$ P - a. e., i.e.

g)

$$P \circ Y^{-1}\left(\left\{ s : \sum_{i=1}^{\infty} <Ce_i, r_j> \lambda_i^{-\frac{1}{2}} <s, e_i> \; < \infty \right\} \right) = 1$$

Namely,

$EY_i = 0, EY_i^2 = 1$ $(i \in \mathbb{N})$ and thus

$$\sum_{i=1}^{\infty} E\left(<Ce_i, r_j> Y_i \right)^2 = \sum_{i=1}^{\infty} <e_i, C^* r_j>^2 \leq \| r_j \|^2 = 1$$

and a) proves g).

Furthermore, for all $j \in \mathbb{N}$: $E\zeta_j^2 \leq 1$. Namely, define

$$\zeta_{j,n} := \sum_{i=1}^{n} <Ce_i, r_j> Y_i$$

and observe Y_i independent and

$$E\zeta_{j,n}^2 = \sum_{i=1}^{n} <Ce_i, r_j>^2 \leq 1$$

Thus, by g), $\zeta_{j,n}$ P - a. e. convergent for $n \longrightarrow \infty$, and also $\zeta_{j,n}^2$ is a P - a. e. convergent sequence of r.v. such that

$$\lim_n \zeta_{j,n}^2 = \zeta_j^2.$$

Thus by Fatou's lemma,

$$E\zeta_j^2 \leq \liminf_n E\zeta_{j,n}^2 = \sum_{i=1}^{\infty} <e_i, C^* r_j>^2 = \| C^* r_j \|^2 \leq 1$$

Eventually, we obtain

$$\| LY \|^2 = \sum_{j=1}^{\infty} \mu_j \zeta_j^2$$

converges P - a. e., since

$$\sum_{j=1}^{\infty} E\left| \mu_j \zeta_j^2 \right| \leq \sum_{j=1}^{\infty} \mu_j = \mathrm{Tr}\, R < \infty$$

and by b),

$$\sum_{j=1}^{\infty} \mu_j \zeta_j^2 < \infty \quad P - \text{a. e., i.e. } P \circ Y^{-1}(D(L)) = 1.$$

Now the estimate is set $x^* := Ly$ for $y \in D(L)$). To show assertion iii) of the proposition, it is remarked, that (X, Z) is a zero-mean Gaussian random element with values in $H_1 \times H_2$ and covariance operator

$$\begin{pmatrix} R & O \\ O & S \end{pmatrix}.$$

Therefore, by the independence of X and Z,

$$(X, Y) = \begin{pmatrix} I & O \\ A & I \end{pmatrix}\begin{pmatrix} X \\ Z \end{pmatrix},$$

and (X,Y) has covariance operator

$$\begin{pmatrix} R & RA^* \\ AR & ARA^*+S \end{pmatrix} = \begin{pmatrix} I & O \\ A & I \end{pmatrix}\begin{pmatrix} R & O \\ O & S \end{pmatrix}\begin{pmatrix} I & A^* \\ O & I \end{pmatrix}$$

Consider the error $\zeta := X - LY$.

<u>It is not immediate that ζ is Gaussian, since L is unbounded.</u>

Let $\Phi_{(\zeta,Y)}((h_1, h_2)) := E_{(\zeta,Y)}\exp[i<(\zeta,Y),(h_1,h_2)>] =$

$$E_{(X,Y)}\exp\left[i<\begin{pmatrix} I & -L \\ O & I \end{pmatrix}(x,y),(h_1,h_2)>\right] =$$

$E_{(X,Y)}\exp[i<X, h_1> - <LY, h_1> + <Y, h_2>]$

be the characteristic function of (ζ,Y).

Let L_{ij} denote the partial sums in the definition of L:

$L_{ij} := P_{H_1^j} L_0 P_{H_2^i}$, where $H_1^j = sp\{r_1, ..., r_j\}$, $H_2^i = sp\{e_1, ..., e_i\}$, and $P_{H_k^j}$ the orth. projection from H_k onto H_k^j $(k = 1, 2 ; i,j \in \mathbb{N})$.

Then

h) $$L_{ij}(ARA^*+S) = P_{H_1^j} RA^* P_{H_2^i}$$

k) $$L_{ij}AR = P_{H_1^j} R^{\frac{1}{2}} C P_{H_2^i} C^* R^{\frac{1}{2}}.$$

Calculating $\Phi^{ij}_{(\zeta,Y)}$, denoting analogously the char. function corresponding to L_{ij}, one gets

l)

$$\Phi_{(\zeta,Y)}(h_1, h_2) = \lim_i \lim_j \Phi^{ij}_{(\zeta,Y)}(h_1, h_2)$$

by continuity of the norm and the exponential function and by Lebesgue's theorem.

m)

$$\Phi_{(\zeta,Y)}(h_1,h_2) = E_{(X,Y)} \exp\left[i < (X,Y), \begin{pmatrix} I & O \\ -L_0^* & I \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} > \right] =$$

$$\exp\left[-\tfrac{1}{2} < \begin{pmatrix} R & RA^* \\ AR & ARA^*+S \end{pmatrix} \begin{pmatrix} I & O \\ -L_0^* & I \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \end{pmatrix}, \begin{pmatrix} I & O \\ -L_0^* & I \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} > \right] =_{\text{by h) and k)}}$$

$$\exp\left\{ -\tfrac{1}{2} \left[<(R - P_{H_1'} R^{\frac{1}{2}} C P_{H_2'} C^* R^{\frac{1}{2}}) h_1, h_1 > + 2 < (-RA^* + P_{H_1'} RA^* P_{H_2'}) h_2, h_1 > + <(ARA^* + S) h_2, h_2 > \right] \right\}$$

By l) one gets

n)
$$\Phi_{(\zeta,Y)}(h_1,h_2) = \exp\{-\tfrac{1}{2}[<(R - R^{\frac{1}{2}}C^{**}C^*R^{\frac{1}{2}})h_1,h_1> +$$
$$(ARA^*+S)h_2,h_2>]\} =$$
$$= \exp\{-\tfrac{1}{2}<D(h_1,h_2),(h_1,h_2)>\} \text{ where}$$
$$D = \begin{pmatrix} R - R^{\frac{1}{2}}C^{**}C^*R^{\frac{1}{2}} & O \\ O & ARA^*+S \end{pmatrix}$$

is again nuclear and non-negativ in $H_1 \times H_2$.

Thus (ζ, Y) is Gaussian in $H_1 \times H_2$, zero-mean, with covariance D.
The form of D gives independence of ζ and Y.
Hence ζ is zero-mean Gaussian with covariance $R - R^{\frac{1}{2}}C^{**}C^*R^{\frac{1}{2}}$
Eventually, $X = \zeta + LY$ and the conditional distribution of X given $Y = y$, $y \in D(L)$, is obtained by
$P(X \in B \mid Y = y) = P(\zeta + LY \in B \mid Y = y) = P(\zeta \in B - LY \mid Y = y)$
$= P(\zeta \in B - Ly) = P(\zeta + Ly \in B) \quad (B \in \mathcal{L}(H_1))$,
i.e. $x^* = Ly$ is the mean of the conditional distribution of X given
$Y = y$, $y \in D(L)$, as in the finite dimensional case.

#

Remarks:

i) L is a measurable linear estimator which dominates the class of continuous linear estimators $T:H_2 \longrightarrow H_1$ without necessarily being a member of this class.

ii) Also in the sense of (4), L gives the best mean square estimate for linear functionals $<X,h>$ of the signal X ($h \in H_1$ arbitrarily fixed).

iii) If X,Z are not centered, we find analogously to the finite dimensional case $x^* = E(X) + L(y-A(E(X))-E(Z))$ $(y \in D(L))$ as LLS-estimate for x with respect to (1) and (4). But additionally we have to require $AE(X)+E(Z) \in D(L)$ to insure that the right side above is well defined.

iv) If the correlation K of Y has a non-trivial null-space, then one considers $H_2 \ominus Ker(K)$. It holds $P \circ Y^{-1}(H_2 \ominus Ker(K)) = 1$. Thus, we have a unique left-inverse for the restriction of K onto H_2, and obtain a measurable linear extension L for L_0, where $L_0 = Q \circ (K|H_2 \ominus Ker(K))^{-1}$, by restricting the extension procedure in Proposition 1 to those eigenvalues of K which are greater than zero. Hence, the above results can be transferred also to this case.

v) Let signal X and noise Z be zero-mean Gaussian with nuclear covariances R and S resp., but correlated with cross-correlation
$Q : H_2 \longrightarrow H_1$. If a joint Gaussian distribution of (X,Z) with correlation
$$\begin{pmatrix} R & Q \\ Q^* & S \end{pmatrix}$$
is assumed, then L_0 corresponds to
$L_1 := (RA^* + Q)(ARA^* + AQ + Q^*A^* + S)^{-1}$.
Again $P \circ Y^{-1}(D(L_1)) = 0$. Analogously to Prop. 1, an extension of the estimation operator L_1 to a measurable transformation L is possible under the additional assumption that $rgQ \subset rgR^{\frac{1}{2}}$.
Then again the estimation error $\zeta = X-LY$ is zero-mean Gaussian, and L is the optimal estimation operator with respect to (1) and (4).

2.2 A finitely additive cylinder measure as noise model

2.2.1 Weak random variables

We make the following assumptions:

i) The signal X is zero-mean Gaussian with values in H_1 and nuclear covariance R.

ii) The noise Z is a zero-mean Gaussian weak random variable with bounded, strictly positive covariance $S: H_2 \longrightarrow H_2$, i.e. a zero-mean Gaussian cylinder measure with covariance operator S is associated to Z.

iii) The signal and the noise are independent.

We look again for an estimate of x in the problem $Ax + z = y$, where A is onto, but not continuously invertible, and $\mathrm{Ker}(ARA^* + S) = \{0\}$ for convenience.
As it is shown in [1], the following result holds:

> If the class of admissible estimators is restricted to Hilbert-Schmidt operators in (1) and (4), then $L_0 y := RA^*(ARA^* + S)^{-1} y$ is the optimal estimate for x in the problem $Ax + z = y$ with respect to the modified criteria (1) and (4), and L_0 is Hilbert-Schmidt itself.

The proof is the same as in section 1., if the restriction to Hilbert-Schmidt operators as admissible estimators is made.

Remarks:

i) L_0y cannot directly be interpreted as mean of the conditional distribution of X given y, since the 'conditional distribution' of X given an observation, which is associated to a cylinder measure on H_2, does not exist in the countably additive sense.
Only for linear functionals of the estimate this term is appropriate in the sense of modified (4).

ii) The restriction to zero-mean random elements is not essential (cf. 2.1). The assumption of independence of signal and noise may be replaced by the assumption that the correlation $K = ARA^* + S$ of the observation is strictly positive.

iii) If the signal-covariance R is allowed to be non-nuclear, then the error covariance $Tr.(R + LKL^* - LQ^* - QL^*)$ is not defined (Q cross-correlation between signal and observation).
However, if K is continuously invertible, then $L_0 := QK^{-1}$ is the best estimator for linear functionals in the sense of modified (4).

iv) If we assume the signal to be a white noise, i.e. a zero-mean Gaussian weak random variable on H_1 with covariance $\sigma^2 I$, and the noise as well zero-mean white noise on H_2 with covariance $\sigma^2 I$, then $L_0 y = \sigma_0^2 A^*(\sigma_0^2 AA^* + \sigma^2 I)^{-1} y$ is the solution of the Euler equation $(A^*A + \alpha B)x = A^*y$ $(y \in H_2)$, where

$$\alpha = \frac{\sigma^2}{\sigma_0^2} \quad \text{and } B = I$$

That means, that the solution of the Tichonov regularization $\|Ax - y\|^2 + \alpha F(x) = \min!$ for the ill-posed linear problem $Ax = y$ is given by our estimate $L_0 y$ in the filtering problem $Ax + z = y$, if the regularization parameter α and the regularization functional F are chosen to be

$$\alpha = \frac{\sigma^2}{\sigma_0^2}$$

and $F(x) = \|x\|^2$ (cf. [12]).

2.2.2 Radonification of the cylindrical model by the concept of Abstract Wiener Space

In the foregoing section we have seen that the solution of the filtering problem $AX + Z = Y$ for a noise model, which is only a finitely additive cylinder measure on the observation space H_2, gives an estimate L_0Y, which is optimal for linear functionals of the signal, i.e. $E(<X-L_0Y, h>^2) = \min\{E(<X-LY,h>^2) : L : H_2 \longrightarrow H_1$ Hilbert-Schmidt operator$\}$. The observation measure $P \circ Y^{-1}$ in this case also is a finitely additive cylinder measure on H_2. In this framework, for instance in [1] the Kalman-Bucy filter equations are developped as equations on the observation space, related to the finitely additive cylinder measure $P \circ Y^{-1}$. In order to get an interpretation of the estimate as conditional distribution of X given Y, in the usual countably additive sense, the concept of Abstract Wiener space is commonly used. This means, that the observation space H_2 is embedded into a larger space W in such a way that a radonification of the cylinder measure to a countably additive probability measure on the Borel sets of W is possible (cf. [4], [9]):

Definition. (cf. [4])

Let H be a real, separable Hilbert space, and W a real Banach space, and μ a zero-mean Gaussian cylinder measure on H.
Let $i: H \longrightarrow W$ be a continuous injection with dense range in W. The triple (i, H, W) is called an Abstract Wiener space, if the norm $\|.\|_1$ of W is measurable with respect to H and μ in the following sense:

For every $\varepsilon > 0$ there is a finite-dimensional projection P_ε on H such that for every finite-dimensional projection P on H orthogonal to P_ε we have:
$\mu(\{x \in H: \|i(Px)\|_1 > \varepsilon\}) < \varepsilon$

Examples and further references can be found in [9].
To apply this concept to our filtering problem $AX + Z = Y$ we make again the assumptions that X is a zero-mean, Gaussian, H_1-valued random element with nuclear covariance R, Z is zero-mean Gaussian white noise on H_2 with covariance I, and independent of X. Then the weak random variable Z defines a zero-mean Gaussian cylinder measure ν, such that for the corresponding outer measure ν^* holds the well-known fact $\nu^*(H_2) = 0$ (cf. [5]). The cylinder measure ν does not have a countably additive extension to the Borel sets $\mathcal{B}(H_2)$ of H_2, since in a certain sense H_2 is too small.

Thus, one extends H_2:

<u>Definition.</u> Let $C:H_2 \longrightarrow H_2$ be an arbitrary linear operator which is

i) self-adjoint, positive semi-definite

ii) injective

iii) Hilbert-Schmidt

Define $<x, y>_1 := <Cx, Cy>$ $(x, y \in H_2)$, $\|.\|_1$ the corresponding norm, and let W be the completion of H with respect to the norm $\|.\|_1$

It is shown in [4], that $\|.\|_1$ is a measurable norm with respect to ν, the weak distribution corresponding to Z. Thus the triple (i, H_2, W), $i:H_2 \longrightarrow W$ the canonical injection, is an Abstract Wiener space. The Gaussian cylinder measure ν induces a Gaussian cylinder measure ν_C on the cylinder sets of W by
$\nu_C(\{w \in W:(w_1'(w), \ldots, w_n'(w)) \in B\}) := \nu(\{h \in H_2:(w_1'(ih), \ldots, w_n'(ih)) \in B\})$ for every choice of $w_1', \ldots, w_n' \in W'$, $n \in \mathbb{N}$, $B \in \mathfrak{L}(\mathbb{R}^n)$.
According to [4], ν_C possesses a countably additive extension ν_C' to the σ-algebra generated by the cylinder sets of W. Since H_2 is separable, this σ-algebra coincides with the Borel-σ-algebra on W (cf. [10]).
Again, ν_C' is zero-mean Gaussian.
We denote the correlation operator of this measure ν_C' by G_1. For the restriction $G_1|H_2$ of G_1 onto H_2 (identified with the subspace $i(H_2)$ in W) we have:

<u>Lemma 1.</u> $G_1|H_2 = C^2$

<u>Proof.</u> For $h_1, h_2 \in H_2$, the function $<.,h_1>_1<.,h_2>_1 : W \longrightarrow \mathbb{R}$ is a cylindrical function. Hence, by definition of ν_C',

$$<h_1, G_1h_2>_1 = \int_W <w, h_1>_1 <w, h_2>_1 d\nu_C'(w) = \int_{H_2} <h, h_1>_1 <h, h_2>_1 d\nu(h) =$$

$$\int_{H_2} <h, C^2h_1> <h, C^2h_2> d\nu(h) = <C^2h_1, C^2h_2> = <h_1, C^2h_2>_1$$

#

Now, we consider the equation $AX+Z=Y$. We transfer canonically the probability measure $P\circ(AX)^{-1}$ onto $\mathcal{B}(W)$:

$$\rho := P\circ(AX)^{-1}$$

$$\rho' := \rho\circ i^{-1} \quad \text{, i.e. } \rho'(B) = \rho(\{h\in H_2 : i(h)\in B\}) \quad (B\in\mathcal{B}(W))$$

ρ' is zero-mean Gaussian and for its correlation operator $G_2: W \longrightarrow W$ holds:

<u>Lemma2.</u> $G_2|H_2 = ARA^*C^2$

<u>Proof.</u> For $h_1, h_2 \in H_2$,

$$\langle h_1, G_2 h_2\rangle_1 = \int_W \langle w, h_1\rangle_1 \langle w, h_2\rangle_1 \; d\rho'(w) =$$

$$\int_{H_2} \langle h, C^2 h_1\rangle \langle h, C^2 h_2\rangle \, d\rho(h) = \langle C^2 h_1, ARA^* C^2 h_2\rangle = \langle h_1, ARA^* C^2 h_2\rangle_1$$

#

The sum of the Gaussian random variables on W, corresponding to ν_C' and ρ', has a measure μ' as its distribution. μ' is also zero-mean Gaussian and has $G_3 := G_1 + G_2$ as correlation operator, due to the assumed independence of ν and ρ and therefore of ν_C' and ρ'. We now have immediately:

<u>Lemma 3.</u>

i) G_3 is nuclear

ii) $G_3|H_2 = (ARA^* + I)C^2$

iii) μ' is the radonification of the cylinder measure $\mu\circ i^{-1}$, where μ denotes the weak distribution of Y in H_2.

For the cross-correlation $G_4: W \longrightarrow H_1$ between the signal and the observation measure holds

<u>Lemma 4.</u> $G_4|H_2 = RA^*C^2$

<u>Proof.</u> For $h_1 \in H_1$, $h_2 \in H_2$,

$$<h_1, G_4 h_2> = \int_{H_1 \times W} <h, h_1><w, h_2>_1 \; P \circ X^{-1}(dh) \mu'(dw) =$$

$$\int_{H_1 \times H_2} <h, h_1><h', h_2>_1 \; P \circ X^{-1}(dh) \mu(dh') =$$

$$\int_{H_1 \times H_2} <h, h_1><h', C^2 h_2> \; P \circ X^{-1}(dh) \mu(dh') = <h_1, RA^*C^2 h_2>$$

If we summarize now these observations, we can state that the assumptions of 2.2.1 are satisfied for the transformed model in the larger space W. Analogously to Proposition 1, we obtain an extended estimation operator $L: W \longrightarrow H_1$ which is a measurable transformation , i.e. $\mu'(D(L)) = 1$. Analogously, Lw gives the mean of the conditional distribution of X given $w \in D(L)$. L is the extension of $L_0 := G_4 G_3^{-1}$, where $\mu'(D(L_0)) = \mu'(rgG_3) = 0$ again. L can be given explicitly on the dense subspace $rg((ARA^* + I)C^2)$ of W:

<u>Proposition 2</u> L is given on $rg((ARA^* + I)C^2)$ by

$$L \mid rg((ARA^* + I)C^2) = RA^*(ARA^* + I)^{-1}$$

<u>Proof.</u>

$$L \mid rg((ARA^* + I)C^2) = G_4 \mid H_2 \circ G_3^{-1} \mid rg((ARA^* + I)C^2)$$
$$= RA^*C^2((ARA^* + I)C^2)^{-1} = RA^*(ARA^* + I)^{-1}$$

by Lemma 3 and Lemma 4.

Thus, a solution of the above linear filtering problem is given in a countably additive framework by using the concept of radonification of the cylinder measures in Abstract Wiener space.

Remark that the solution is independent of the renorming operator C by which the Abstract Wiener space was constructed. It is formally equal to the 'cylinder measure solution' and differs from it only by its domain of definition, which makes it a measurable transformation from W to H_1 with respect to μ'. Hence, the martingale approach to the Kalman-Bucy filter in the countably additive theory, as worked out in [7], gives formally the same result as the 'cylinder measure filter' in [1]. We conclude with an argument that shows the significance of the finitely additive filtering theory, recently worked out by G. Kallianpur and R. L. Karandikar [7], namely that the space H_2 of the actual observations has zero probabiltiy with respect to the radonified observation measure μ' in W.

Proposition 2 With the above notations, identifying H_2 with $i(H_2)$, it holds: $\mu'(H_2) = 0$

Proof: As is well-known, W is isometrically isomorphic to a closed linear subspace F of the space C[0,1] of continuous functions on the unit interval. If this isomorphism is denoted by ψ:W$\longrightarrow$F, and H denotes the reproducing Kernel Hilbert space generated by the covariance of the Gaussian measure $\nu_C' \circ \psi^{-1}$, then $H_2 = \psi^{-1}(H)$, according to [6]. If B is the covariance of $\nu_C' \circ \psi^{-1}$, then $x \in H$ iff $x \in rg(B^{\frac{1}{2}})$, due to [3]. Thus $\nu_C'(H_2) = \nu_C' \circ \psi^{-1}(rg(B^{\frac{1}{2}})) = 0$.

The last equality holds analogously to (5).

Eventually, using a result of G. Kallianpur and R.L. Karandikar on mixtures of translates of the canonical Gauss measure (cf. [8]), we have

$$\mu'(H_2) = \int_{H_2} \nu_C'(H_2 - h)\rho(dh) = 0$$

#

If Z is a white noise in the model AX + Z = Y on an observation Hilbert space H_2 (often of smooth functions; for example H_2 the RKHS of the Brownian motion), this precisely means, that the filter solution with respect to the radonified observation measure μ' on Abstract Wiener space W does not have a practical meaning as long as observations are considered to be elements of the space H_2 (cf. [8]), i.e. as long as they appear with zero probability.

Acknowledgements

I would like to thank Professor G. Kallianpur for his kind invitation to the Center for Stochastic Processes. I am greatly indebted to him and also to Professors S. Cambanis and R. Leadbetter for their interest and fruitful discussions on the subject of this work.

References

[1] Balakrishnan, A. V. Applied Functional Analysis Applications of Math. 3, Springer 1981

[2] Franklin, J. N. Well-Posed Stochastic Extensions of Ill-Posed Linear Problems J. Math. Analysis and Appl. 31, 682-716 (1970)

[3] Gihman, I. I. and Skorohod, A. V. The Theory of Stochastic Processes Springer 1974.

[4] Gross, L. Abstract Wiener Measure and Infinite Dimensional Potential Theory Lect. Notes in Math. 140 (1970), Springer

[5] Kallianpur, G. The Role of Reproducing Kernel Hilbert Spaces in the Study of Gaussian Processes Advances in Probability and Related Topics, II, ed. by P. Ney, M. Dekker Inc., 49-83, N.Y. (1970)

[6] Kallianpur, G. Abstract Wiener Processes and Their Reproducing Kernel Hilbert Space Z. Wahrscheinlichkeitstheorie 17, 113-123 (1971)

[7] Kallianpur, G. Stochastic Filtering Theory Appl. of Math. 13, Springer 1980

[8] Kallianpur, G. and Karandikar, R. L. White Noise Calculus and Nonlinear Filtering Theory
Ann. Probability 13 (1985)

[9] Kolzow, D. A Survey of Abstract Wiener Space
Proc. Summer Res. Instit. on Statistical Inference for Stochastic Processes.
Indiana Univ., Bloomington, Ind., Vol. 1, 293-315 (1974)

[10] Sato, H. Gaussian Measures on a Banach Space and Abstract Wiener Space
Nagoya Math. J. 36, 65-81 (1969).

[11] Strand, O. N. and Westwater, E. R. Statistical Estimation of the Numerical solution of a Fredholm Integral Equation of the First Kind
J. Assoc. Comput Mach. 15, 100-114 (1968)

[12] Tichonov, A. N. and Arsenin, V. Ya. Methods of Solving Ill-Posed Problems
Nauka, Moscow 1974

[13] Uhlig, A. Untersuchungen zur stochastischen Regularisierung fur lineare Gleichungen in Hilberträumen
Dissertation, TH Karl-Marx-Stadt, 1977

END
DATE
FILMED
DEC.
1987